

## ENTREVISTA A ISIDRO F. AGUILLO

### **EL VERDADERO PROBLEMA DE LOS MOTORES DE B SQUEDA DE INTERNET NO ES EL RUIDO DOCUMENTAL, SINO EL SILENCIO**

participado en varios proyectos del V Programa Marco de I+D de la Unión Europea y del Plan Nacional de Investigación Científica y actualmente sigue trabajando en parte de ellos.

Siempre en relación con las últimas tecnologías, Isidro Aguillo ha sido miembro de la Oficina Española de Ciencia y Tecnología (SOST) en Bruselas, Metcalfe Visitor Professor en la Universidad de Nueva Gales del Sur (Sidney, Australia) y en la actualidad es miembro de la Junta Directiva de la Sociedad Española de Documentación (SEDIC). Desde 1997 dirige

En 1996 Aguillo recibió el premio IWE 96 por votación de los documentalistas españoles como "Personalidad del año".

-¿De qué se ocupa el Departamento de Innovación Tecnológica del CSIC y cuáles son las funciones de usted en él?

-El Departamento lo constituimos varios equipos con misiones diferentes, pero complementarias. Hay un equipo que se dedica a desarrollos informáticos, a poner productos que puedan ser utilizados por



## VISTA ENTRE

Isidro F. Aguillo

Lourdes Toledo



"La digitalización ya se está llevando a cabo, y con éxito, pero se puede y se debe dar pasos más adelante"

Isidro F. Aguillo coordina actualmente el Laboratorio de Internet en el Centro Nacional de Información y Documentación Científica (CINDOC) del Consejo Superior de Investigaciones Científicas (CSIC). Con una trayectoria profesional brillante, Aguillo, licenciado en Biología por la Universidad Complutense y Máster en Información y Documentación por la Universidad Carlos III de Madrid, es hoy por hoy un especialista de renombre en temas relacionados con la Sociedad de la Información. Como tal, trabaja en el Departamento de Innovación Tecnológica del CINDOC, donde realiza tareas relacionadas con el desarrollo de indicadores de la Sociedad de la Información, el análisis documental de recursos Web, la Cibermetría (valoración cuantitativa de la información que aparece en la Web y de sus flujos de circulación) y procesos de comunicación científica a través de la Red.

En los últimos años ha dirigido y ha

la revista electrónica "Cybermetrics" y es miembro del Comité Asesor del "Profesional de la Información" y de comités científicos de diversos congresos nacionales e internacionales. Además, participa como evaluador y revisor de proyectos europeos de investigación.

En el terreno académico, ha impartido desde 1994 más de trescientos cursos y seminarios en cuarenta y cinco universidades españolas y extranjeras, así como en doce asociaciones profesionales y veintiocho instituciones, tanto públicas como privadas. Desde hace tiempo su agenda está ocupada por numerosos cursos y seminarios que solicitan su presencia. El pasado mes de octubre Isidro Aguillo estuvo en Valencia, donde impartió un curso organizado por el Instituto Valenciano de Administraciones Públicas (IVAP) y la Dirección General del Libro y Bibliotecas, bajo el título "Recuperación de Información para bibliotecarios, documentalistas y archiveros"

intermediarios, documentalistas y bibliotecarios, entre otros, para diferentes aplicaciones documentales. En este sentido, últimamente nos hemos dedicado mucho a los Tesoros y hemos puesto en el Web algunos de los que teníamos ya diseñados de manera que sean dinámicos e interactivos y que se pueda navegar por ellos pasando de un término a otro, algunos de ellos relacionados, y siempre de manera muy ágil. En Web esto es fundamental, ya que estamos hablando de ficheros considerablemente grandes. Un segundo equipo, que se encarga de cuestiones terminológicas, está centrado en proyectos compartidos con la Unión Europea. A Europa le interesa disponer de servicios multilingües y en este terreno nosotros somos competentes por nuestra experiencia en desarrollo terminológico en diferentes idiomas. Un tercer grupo se dedica al desarrollo de portales de Internet y trabaja normalmente bajo encargo, exceptuando el caso del portal de ciencia y tecnología,

considerado de interés prioritario para el Consejo. Hasta la fecha han creado portales sobre temas con un fuerte componente social, como son los trastornos de alimentación: (anorexia y bulimia), o un portal sobre el agua que estamos creando para el CEDEX, perteneciente a Fomento. Aparte, está también el grupo de trabajo que yo coordino "Laboratorio de Internet", dedicado a desarrollos generales sobre análisis documental en Internet, en especial en términos cuantitativos, como es la presencia en el Web, fundamentalmente en el sector académico y de investigación. Dentro del

*Hay una necesidad de intermediarios entre los usuarios y la informacin. Y ah tenemos un reto los profesionales de la informacin. La cuestin clave es si vamos a dejar que unos colectivos que no estn tan preparados como nosotros ocupen esa posicin en nuestro lugar*

mismo equipo, llevamos también algunos proyectos europeos, mediante los que podemos conocer el desarrollo de Internet en diferentes países y observar, no lo que se ha llamado la brecha digital, pues estamos trabajando en un entorno que ya conocemos, sino los flujos de la información y sus diferentes recorridos con un método supuestamente objetivo como es el análisis cuantitativo.

-Estamos hablando de Cibermetría. ¿No es así?

- Sí, bueno. El término es un poco variable porque fuera de España se llama webometrics, pero webometría sonaba extraño en castellano y la palabra que se decidió usar fue cibermetría.

-¿Cuál es su visión sobre el factor de Impacto Web en las instituciones académicas de nueva creación, que, a pesar de tener poca producción científica, han invertido mucho en su presencia en la Red, como

es el caso de la Universitat Jaume I de Castelló?

-El fenómeno de la Jaume I es admirable, creo que fueron ellos los que crearon el segundo Web del Estado español. En cambio, si lo analizamos desde un punto de vista un poco más amplio, no sólo debemos fijarnos en el factor de Impacto Web, que sólo es uno más, sino que hay que medir y evaluar la presencia en el Web. El Web es cada día más prioritario, y por ello estar presente en él genera visibilidad e interés, además de incrementar la trascendencia institucional. Hasta la fecha se ha cometido el error de fijarnos sobre todo en los datos que llamamos físicos como el número de ordenadores conectados a la Red y el de servidores, y se ha prestado muy poca atención a la valoración y la evaluación de los contenidos que transporta Internet. Por otra parte, te encuentras con muchas sorpresas, y es que todavía hay instituciones con tradición y un peso importante que siguen fuera de la Red porque no han asumido la prioridad que ésta ya es.

-¿Qué métodos y qué criterios utilizan para medir la presencia en el Red, así como los movimientos de los flujos informativos dentro de ésta?

-Lo que hacemos es valorar el impacto, no sólo general, sino específico. Por ejemplo, la influencia que puedan tener los centros académicos del Estado Español sobre otras instituciones europeas o iberoamericanas, y aquí son notables las carencias existentes para un país que pretende reforzar la presencia del castellano en la Red. El impacto de las instituciones académicas españolas en las mejicanas, argentinas, colombianas o venezolanas es sorprendentemente limitado. Muy escaso, diría yo. Respecto a Europa, la situación cambia un poco y parece que se esté logrando cierto éxito, de manera que la cooperación intracomunitaria tiene ya un reflejo razonable en los Web, pero en términos generales, las sedes Web españolas son pobres en información y se están valorando, en la mayoría de los casos, siguiendo criterios de formato como la elegancia, el atractivo y la facilidad de navegación, que si bien son importantes, a mi modo de entender lo son mucho más los contenidos. Todo ello me lleva a afirmar que estamos en una situación muy lamentable.

-¿Cómo podemos mejorar la recuperación de la información y conseguir que ésta sea

fiable y pertinente para las necesidades que teníamos antes del comienzo de la búsqueda?

-Lo primero que quiero decir es que la actual batería de motores de búsqueda tiene carencias notables, a pesar de lo que en general se piensa de ellos. Es el caso de Google y lo que yo llamo el "googlecentrismo", que considero tiene carencias porque hay información que nosotros los documentalistas estamos acostumbrados a buscar en bases de datos y que no está disponible en los motores de búsqueda y que sería importante que estuviese. Ahora bien, lo más grave es la actitud de la sociedad hacia los motores de búsqueda. En general, se piensa que buscar en Internet es tan simple como poner un término en cualquiera de los motores de búsqueda y darle al enter, lo cual es un error, que además, ha trascendido ya en el ámbito bibliotecario y documentalista y ha causado que mucha gente no sepa optimizar los recursos que le ofrece Internet y piense que se genera mucho ruido en la Red y que la calidad de los resultados no son buenos. Considero que lo que se está ocultando, en la mayoría de los casos, es una incorrecta utilización de los resultados de las búsquedas. A mi modo de ver, el verdadero problema de los motores de búsqueda no es el ruido, sino el silencio documental. El ruido puede ser filtrado, pero si se genera silencio, no hay una herramienta adecuada para solventar el problema, por eso creo que cuando se afirma que la precisión y la tasa de resultados de búsqueda positivos no es muy elevada, en la mayoría de los casos estas afirmaciones ocultan un desconocimiento profundo sobre los métodos de recuperación en Internet, que son diferentes de los métodos de las bases de datos, donde el control documental que tienen éstas no es por ahora el de Internet. Todo ello ha dado lugar a algunas posiciones que apuntan hacia la necesidad de dotar a Internet de un control documental, para solucionar el carácter caótico -dicen- que tiene la Red. Para mí ése no es el problema, sino que hay que respetar de una manera democrática lo que cada cual quiera poner en Internet y no hay que forzar a nadie. Por otra parte, proyectos como el de reorganizar toda la información en Internet serían muy difíciles de implementar. Pienso que hay que tener

en cuenta esa diversidad en la Red, y sobre ella, aprender cómo funciona, sin olvidar que el control documental de Internet, insisto, es menor del que existe en las bases de datos, y aplicarlo de manera coherente a la hora de recuperar la información. Una vez hecho esto, hay que cambiar algunos aspectos. Por ejemplo, hoy por hoy se está enseñando a buscar en Internet como si lo que pretendiéramos fuera resolver preguntas de concurso de televisión, lejos en la mayoría de los casos del objeto de la búsqueda. Además, se está dando por

de una sede Web?

-Creo que la mayoría de las propuestas que se han presentado hasta ahora son propuestas formales, e incluso algunas que mezclan características formales y substanciales, de contenido. El problema es que al mezclar las dos consideraciones hay información valiosísima que se diluye. Por ejemplo, en cualquier obra publicada es muy relevante el reconocimiento de la autoría, en cambio, en Internet no es un factor tan decisivo a la hora de considerar si la información es o no fiable, lo cual me parece lógico, ahora bien, la autoría institucional

la sede es hasta qué punto es o no navegable...Pues ahí entra la duda. -¿Qué ocurre con el diseño de muchas páginas web que a veces no nos ofrecen información útil como puede ser una ficha técnica que nos facilite la comunicación con el organismo o la persona que representa esa sede?

-El problema creo que se debe a la ausencia de retroalimentación. Hay que hacer entender a los Webmasters que esa información es relevante, y que si ellos la incluyen, facilitan el acceso a los recursos de los sitios en cuestión, lo cual es ventajoso respecto a los Webs que no proceden así. Todos esos datos se han de tener en cuenta en el posicionamiento de un motor de búsqueda, en un índice, por ejemplo, ya que supone un valor añadido. ¿Por qué esto no ocurre así? Piensen que el índice que más se está utilizando hoy por hoy es Yahoo, "el modelo Yahoo" que llamamos nosotros, aunque en realidad ocurre lo mismo con muchos otros índices. Sucede que cuando uno entra en Yahoo, la descripción que encuentra de una página es ridícula, por ejemplo: "Esta es la página de x y poco más" y es algo totalmente absurdo, que no exista ningún tipo de valoración o de descripción en estos niveles, lo que hará que otros Webmasters no le presten atención.

-¿Qué análisis documental sería entonces el más eficaz y qué criterios los más apropiados para evaluar la calidad de las sedes Web?

-Nosotros lo que proponemos es que se realice un análisis documental más extenso, de tipo descriptivo, por ejemplo: cuál es el volumen de información que se ofrece, si se ofrecen mecanismos de navegación, si en la página principal hay un índice, un mapa o un buscador interno de la sede, si la autoría figura o no de manera explícita, (la cual se puede describir a un nivel muy profundo, ya que podemos exigir no sólo el nombre del autor, sino una dirección electrónica o postal de contacto), así como la opinión que ofrecen sobre esa serie de páginas otros expertos (con los que también debe poderse contactar). Ese tipo de información es lo que ahora en España se llaman "directorios analíticos" y ya hay algunos ejemplos llevados a la práctica, como el de "Darwin", un proyecto de varias universidades españolas, liderado por la Universidad de Salamanca. La idea básica es que si el Webmaster encuentra algún intermediario que ha descrito características

hecho que se puede resolver una recuperación de información en Internet recurriendo sólo a dos, tres, cuatro o cinco sitios de la primera o segunda página de respuestas, y el tratamiento que hay que hacer en la Red debe ser más exhaustivo. Es decir, que si estoy recuperando información, tengo que tener en cuenta todas las posibles visiones, porque la información que ofrece Internet ya no es monolítica sino que responde a diferentes idiosincrasias, diferentes culturas, idiomas e instituciones. Resulta absurdo, por tanto, planteamos únicamente la revisión de las cuatro o cinco primeras respuestas de una búsqueda, hay que considerar un nivel viable, que podría estar en los pocos centenares, y para eso existen herramientas automáticas que nos permiten recuperar estos recursos de una manera rápida y directa y luego proceder a analizarlos.

-¿Cómo se puede analizar la fiabilidad

es posible reconocerla a través de mecanismos indirectos. Otro de los puntos que se tiene muy en cuenta es la navegabilidad, pero se olvidan otros aspectos como el tamaño de las sedes Web. Hoy por hoy, nos encontramos con tamaños inusitados, que no tienen nada que ver con lo que habíamos visto hasta ahora, y si bien es importante desde la óptica de la navegación, plantea unos problemas importantes a la hora de la descripción. Tradicionalmente, hemos podido describir un objeto documental con dos, tres o cinco párrafos, pero resulta absurdo describir de la misma manera una sede Web con cientos de miles o incluso unos pocos millones de páginas. Estoy pensando en las páginas de la Unión Europea o de la NASA, que tienen varios cientos de miles de páginas. Querer reducir su información a un resumen de unos pocos párrafos supone quitarle mucho valor añadido a la descripción de esa sede. Entonces, si el único criterio que vamos a utilizar para valorar



negativas sobre su sede, intente solventarlas, del mismo modo que se preocupan por el posicionamiento a través de las etiquetas meta de descripción de contenidos, algunas de las cuales identifican al autor de la página, otras contienen las palabras clave que el autor ha elegido porque cree que son las que mejor describen la página, con lo cual la calidad de la descripción suele ser mayor. Estas identificaciones (autor, idiomas en que aparece la página, fecha de última actualización, entre otras) tienen un valor muy importante porque son un sistema precatalográfico que nos ayuda bastante en la descripción de la sede, ya que son valores que nos pueden ayudar a definir la calidad de la sede, al mismo tiempo que hacen aumentar el impacto potencial de la página. ¿Pero nos están poniendo las etiquetas meta para ayudarnos? No, lo hacen simplemente porque saben que usándolas, su posicionamiento en el buscador es mejor. En resumen, existen varios criterios para evaluar la calidad de los Webs, pero si algunos ya vienen dados, como ocurre con las etiquetas meta, la labor de selección es más eficaz, aunque quiero insistir una vez más en que indicar autoría, fecha e idiomas es un aspecto importante, pero poco más que formal. El hecho de que una página no indique estos elementos no quiere decir que el recurso sea negativo, los contenidos pueden ser muy buenos, simplemente no se han tenido en cuenta los aspectos formales, que son valores añadidos.

-¿Deberían entonces los usuarios de Internet no dejarse influir tanto por la estética de la página y profundizar más en los contenidos de las sedes Web?

-Hay una tendencia a crear ránquings y no creo que eso sea realmente necesario. Bastaría con ofrecer una estructura descriptiva, un directorio o una serie de recursos que has seleccionado, de los cuales describes las características. El usuario final puede, posteriormente, tenerlos en cuenta o no, valorarlos según sus necesidades. Y si estamos interesados en que esos recursos sean utilizados en un catálogo de una biblioteca, el hecho de no figurar la autoría nos echaría ya para atrás, mientras que si lo que nos interesa no es la referencia Web bibliográfica, sino los contenidos, el autor ya es secundario.

-¿Qué función tendrían las bibliotecas, los archivos y los centros de documentación a la hora de ofrecer información sobre sus

servicios?

-Hay una función que ya ha sido asumida por el mundo institucional, y es la de representatividad. Es decir, todas las instituciones, independientemente de su carácter público o privado, de que tengan o no ánimo de lucro, deben estar presentes en Internet porque la gente utiliza cada día más este canal para acceder a la información. Creo que todos los centros que trabajan directamente sobre la información deberían profundizar un poco más en los servicios que están ofreciendo. Por ejemplo, el servicio tradicional electrónico de una biblioteca ha

posible mantener permanentemente en una página Web noticias de actualidad. Y sobre todo, la biblioteca pública debería centrarse en ofrecer un portal local, que es algo para lo cual posiblemente está más capacitada que muchas otras instituciones, como son ayuntamientos o colectivos locales. Respecto a otras bibliotecas, como las especializadas, al igual que la biblioteca pública conoce a sus usuarios, éstas deben también conocer a los suyos. Hay una experiencia muy interesante que están llevando a cabo todas las bibliotecas universitarias

*Me preocupa mucho que el bibliotecario o el documentalista iniciara programas de selección, convirtiéndose en una especie de censor. Creo que se puede censurar bastante describiendo, pero dejando siempre la oportunidad al usuario final, quien es el responsable último de seleccionar*

sido el OPAC, pero el OPAC no es la herramienta más adecuada para recuperar información desde el Web porque el objeto último que se recupera es simplemente una referencia bibliográfica, con lo cual el OPAC tiene que estar presente en el Web de una biblioteca pero no debe ser lo que vertebral la sede. La biblioteca ofrece muchos más servicios que debe poner a disposición de sus usuarios a través de la Red. ¿Qué servicios son? Pues los mismos que sirve físicamente, pero a través de la Web. Por ejemplo, las bibliotecas públicas deberían presentar pantallas de navegación especialmente pensadas para sus usuarios, con recursos de referencia, recursos pensados y diseñados para sectores especialmente visitantes de la Web, público infantil, jóvenes, tercera edad, personas discapacitadas, recursos de ocio, como puede ser información sobre el ocio local, noticias, por ejemplo. La sindicación de noticias es muy interesante. Con un mecanismo sencillo y apenas sin costo es

americanas, mediante la que se intenta poner un único punto de entrada a todas sus bases de datos, y por tanto desarrollar unos mecanismos que permitan, no sólo buscar en sus OPACS, sino en todas las bases de datos a las que ellos tienen acceso, así como las que ellos mismos generan. Por otra parte, un caso típico que está empezando a entrar en España, pero que hasta ahora no era frecuente, es el tema de las tesis doctorales a texto completo. Estos sistemas permiten que con una única búsqueda o punto de acceso, podamos acceder simultáneamente a toda la información sobre un tema en concreto que está ofreciendo esa universidad. Por lo que respecta a otro tipo de organizaciones, como son los archivos y los centros de documentación, éstos han de tener en cuenta el riesgo de pérdida con el tiempo de lo que llamamos el patrimonio digital, porque la propia Web tiene una alta tasa



de recambio, es decir, es muy volátil. Podemos preguntarnos por qué, entonces, un archivo que se dedica a guardar documentos relevantes en soporte papel no guarda memoria histórica de las páginas Web cada cierto tiempo, digamos en un ámbito regional, que es una extensión razonable, donde se podría conservar todo lo que diversas instituciones públicas han ido poniendo en la Red en los últimos años e ir guardando la copia íntegra de esos sitios con cierta frecuencia, cada tres o seis meses. Creo que los archivos tienen que ocuparse de preservar

ponerse ya en la Red. El centro de documentación tiene esa misión de aunar todos los recursos de información de la institución, tanto los públicos como los innatos que ella genera.

-¿Qué ocurre en el ámbito académico, donde usted trabaja?

- Aquí el reto fundamental es la comunicación informal, es decir, los científicos fundamentalmente se comunican y valoran su comunicación a través de los artículos científicos publicados en revistas de impacto, pero ese no es todo el conjunto de la actividad que

solicitado, con qué objetivo, los resultados obtenidos de la explotación de esos datos y establecer unas normas de manera que esa información sea realmente utilizada), de manera que puede ofrecer seguridad delante del miedo que a veces puede tener un científico de que alguien acceda a sus datos y los utilice de forma no adecuada.

-¿Cuál es la labor pues del profesional de la Información?

En los tres casos, bibliotecarios, documentalistas y archiveros, creo que su labor ha de ser facilitar la información, pero no la que tradicionalmente han puesto a disposición de los usuarios, sino utilizar una visión más amplia y añadir nuevas tipologías de información, forzar para que los actores que están poniendo información sean más generosos y nos ofrezcan, no sólo información formal, ya finalizada y perfilada, sino también aquella informal que puede ayudar a regenerar los procesos de génesis del conocimiento, los procesos de comunicación interna en una institución, así como ilustrar la información de un cierto producto documental. En definitiva, ir más allá de la labor tradicional. En ese sentido, tengo que decir que la digitalización ya se está llevando a cabo, y con éxito, pero se puede y se debe avanzar más.

-En general existe el sentir que cada vez tenemos acceso a más abundante y variada información, aunque también más desestructurada. ¿Cuál es la responsabilidad de los profesionales de la información (bibliotecarios, archiveros, documentalistas) ante el maremágnum informativo con que se encuentra el visitante de la Red ?

-Me gustaría insistir en que no es malo que exista esa diversidad. En principio, yo no me preocuparía por si el volumen de información es enorme o si el tratamiento posible de esa información es dificultoso porque creo que debe mantenerse así y que cualquier otra cosa se acercaría demasiado al concepto orweliano de "El Gran Hermano", lo que no quiere decir que no debemos asumir ciertas responsabilidades. En realidad, en Internet no hay muchas novedades respecto a la manera como hemos clasificado la información tradicionalmente y cuando acudes a sitios bien clasificados te encuentras con sorpresas. Es el caso de Yahoo, que teóricamente es un modelo creado ex profeso para la Web. En este sentido, hay estudios



"Los últimos datos dicen que aproximadamente el 1,5% de todo Internet tiene un carácter negativo, el cual incluye sectas, pornografía, nazismo, defensa de la anorexia y la bulimia"

ese patrimonio digital, y no sólo transferir a formato electrónico documentos que no estaban en papel, sino conservar lo que se está poniendo en la Red. Todo esto es algo muy novedoso, que todavía no se está llevando a cabo. En cuanto a los centros de documentación, parece ser que éste todavía es un terreno de las intranet, y creo que hay que empezar a abrirse al público. Aquí entramos en el asunto de la gestión del conocimiento. Las instituciones, no sólo las empresas privadas, sino también las públicas, generan una serie de conocimientos intangibles que no están siendo explotados porque los actuales mecanismos para explotarlos, lo que se conoce como gestión de conocimiento, están logrando cierto éxito en entornos internos, en intranets. Ahora bien, alguna de esa información -sobre todo en el sector público, donde los problemas de competencia son menores- debería

desarrollan. Hay que pensar que un científico asiste a un congreso o conferencia y presenta transparencias en formatos Power Point. Por qué éstas no están disponibles en la Red, o dónde se puede encontrar el material con el que un académico da clases en la Universidad, como son transparencias, bases de datos, memorias o informes que desarrollan para terceros, y que ofrecen un análisis más detallado que el del artículo científico. Incluso, en caso extremo, por qué no poner en la Red todos los datos que ha utilizado. Normalmente, por problemas de espacio, ya que en una revista se ofrecen datos generales o un resumen, pero no se da la hoja de cálculo completa. ¿Por qué no colgarla de Internet esa hoja? El problema que puede existir es, primero, si el académico o investigador en cuestión sabe hacerlo, y aquí el documentalista puede establecer una serie de criterios para que le resulte fácil y cómodo, y en segundo lugar, el mismo documentalista podría llevar a cabo un control del acceso y del uso de esa información (qué gente los ha

hechos por OCLC que demuestran que el 95 % de las categorías de Yahoo tienen su equivalencia directa en categorías de la CDU. Si la clasificación bibliotecaria tradicional ha funcionado incluso en un entorno tan desestructurado y caótico como puede ser la producción editorial, ¿por qué no aplicarlo al Web? Pues cuando se ha hecho así, el resultado ha sido muy bueno. Volviendo a la idea del exceso de información, creo que el gran volumen de información que puede suponer un problema para un número restringido de usuarios, puede resolverse. Lo que hoy nos falta son fundamentalmente perspectivas locales y lo que marea a la gente es pensar que toda lo que hay en la Web le interesa. No era sí antes de la Red, por qué tendría que serlo ahora. Antes la gente acudía a las fuentes locales cuando necesitaba recursos sobre un tema y resolvía perfectamente el problema. Hoy, el hecho de saber que además existen otras fuentes lejanas lo único que debería indicarnos es que en caso de necesitarlas están disponibles, pero nada me obliga a tener que utilizarlas. Por lo tanto, la labor es sobre todo previa, de identificación y valoración, pero también de descubrir cuáles son las carencias y las lagunas existentes, como por ejemplo, un caso que me es especialmente doloroso, la escasez que hay en Internet de recursos propios en español y en el resto de las lenguas oficiales del Estado.

-¿No sería una de las principales responsabilidades de los profesionales de la información la de orientar sobre cómo buscar información de calidad?

-Una de las cuestiones que el documentalista puede sacar a la luz mediante una labor específicamente descriptiva son esas lagunas a que me refería antes. A mí me preocupa que nuestros hijos el día de mañana, utilizando Internet, puedan encontrar mucha información sobre la conquista del Oeste norteamericano y que no exista ninguna fuente de información sobre Aníbal o Viriato, o incluso sobre Aristóteles o Platón, y eso ya está ocurriendo. Es en este terreno, creo yo, donde el documentalista debería ser capaz de orientar su programa de evaluación y descripción de recursos para sacar a la luz esas carencias, e incluso, si está en posición de hacerlo, convencer a la gente que tiene esa información para que la incluya en la Red. Esa labor tiene que estar en la agenda de trabajo de los documentalistas, capaces

de hacerse visibles mediante sus recursos a aquellos gestores y políticos que pueden influir en posibles autores para que pongan esa información en la Web. ¿Cómo se puede lograr esto? Con Webs muy útiles, muy listadas y con un alto valor añadido, y con una toma de conciencia al respecto, de manera que cuando aparezca una laguna, se haga pública y alguien se escandalice. Hay que ser muy exhaustivos, descubrir todo lo que hay (y lo que no hay) para ilustrar las carencias. Yo sigo pensando en el modelo de un índice que supere al de Yahoo. Desearía que empezara a haber en España índices descritos de acuerdo con el modelo "Dublin Core", que permite describir una página a partir de quince campos diferentes, que tiene un valor añadido para el bibliotecario porque es precatalográfico y se puede poner como parte de un índice descriptivo, además de ser perfectamente incorporable a un catálogo.

-Por tanto los bibliotecarios, documentalistas y archiveros: ¿deben ser buenos orientadores respecto a la búsqueda y rescate de la información, además de saber seleccionar la que ofrecen ellos?

-Deberían serlo. Me preocupa mucho lo que pueda ocurrir en las bibliotecas escolares, tanto de enseñanza primaria, como de secundaria y el acceso indiscriminado de la juventud, en general, a Internet. Hasta la fecha las bibliotecas escolares no han dispuesto de un profesional específico del ramo, porque se suponía que estas funciones no justificaban quizá este puesto de trabajo, lo que es muy cuestionable, por supuesto. Y en ese sentido, insisto en que en la actualidad, es nula la posibilidad de que exista una persona que de forma activa ayude a que los escolares y estudiantes accedan a la Red y sepan cómo recuperar correctamente la información, siendo capaces, por ellos mismos, de reconocer una serie de características que ilustran cuáles son esos sitios conflictivos. Esta situación justifica, con creces, que un bibliotecario especialista en Internet ocupe ese puesto en las bibliotecas escolares. La respuesta a la pregunta inicial es sí, pero hay que ser mucho más ambiciosos, y aplicar el trabajo de los profesionales de la información al sector público, para también en una serie de instituciones cuya presencia en Internet tendría un valor añadido, y que muchas veces no se tiene en cuenta, como es el caso de las PYMES, que individualmente, una por una, no pueden tener mucha frecuencia, pero se pueden crear consorcios de PYMES por sectores. Un documentalista puede gestionar cosas tan

importantes como hacer comercio electrónico, un tipo de comercio que está creciendo con fuerza hoy por hoy y que todavía la mayoría de las PYMES no tiene acceso a él. En ese caso, creo que los documentalistas podrían intervenir y ayudar en el proceso, no sólo de cara al cliente final, sino ayudar en los pedidos, resolver problemas de importación y papeles de aduanas, entre otras cuestiones. Y volviendo a los centros públicos, insisto en que el documentalista debe ser motivador para que toda la información pública esté en la Red y que no se quede esta intención en un simple *desideratum*, sino que legalmente los centros públicos están obligados a hacerlo. Lo que ocurre a menudo es que posiblemente éstos no saben cómo hacerlo o se plantean problemas de tipo legal, a causa del desconocimiento que hay fuera de nuestro sector respecto a toda la legislación que existe sobre los derechos de autor. Es ahí donde deberíamos trabajar más los profesionales de la información, así como en la explotación de los datos del consumo de información y que el sector público informe a través de la Red sobre cómo está siendo esa información utilizada por los consumidores: quién, con qué fines, por qué. Hoy por hoy el sector público se está olvidando de todo esto.

-¿Cómo se está trabajando actualmente en España respecto al seguimiento de los flujos de la información?

-Casi exclusivamente a nivel descriptivo. En cuanto a las posibilidades analíticas que pueden mostrarnos cómo llega el usuario a la información, qué ruta sigue, por dónde sale, qué páginas ha visitado, cuánto tiempo se ha detenido en cada una, si ha hecho algún tipo de descarga, dónde la ha hecho, dónde le ha llevado la página, si ha vuelto a ella después de salir, hay muy poco hecho. De manera que si casi es inédito el análisis descriptivo, el analítico, es realmente extraño que se presente

-¿Tendría alguien que encargarse de seleccionar qué información es útil y cuál no, o quizá no sea esta la mejor manera de orientar al usuario sobre el contenido? ¿Cree usted que este servicio se está ofreciendo actualmente a los usuarios?

-Entre el usuario y la información hay una necesidad de intermediarios, y ahí

tenemos un reto. La cuestión clave es si vamos a dejar que unos colectivos que no están tan preparados como nosotros ocupen esa posición en nuestro lugar, como ha ocurrido hasta ahora con los informáticos. Creo que la decisión última está en nuestras manos y no en la de ellos, su labor es claramente otra y en ese sentido, deben estar apoyando y ayudando a desarrollar esas herramientas automáticas que son necesarias, pero cuando se trata de gestionar la información, tenemos que ser nosotros los que intervengamos. Si el usuario se encuentra desorientado, se verá

buscador sea eficaz, y no la seguridad. El informático debe ser capaz de traducir esos requerimientos en términos informáticos, pero lo que jamás entenderé es que me pongan como prioritario el tema informático, no sólo a mí, sino a mi equipo, a mi institución y al tipo de servicios que presto. Creo que cuando hayamos logrado evitar esto, la colaboración con los informáticos será maravillosa y la sinergia será mejor. El informático, insisto, es un técnico de apoyo y las prioridades las decide la gente que tiene responsabilidades, los bibliotecarios, los documentalistas o los archiveros.

pero dejando siempre la oportunidad al usuario final, quien es el responsable último de seleccionar. Y respecto a este tema, me gustaría volver rápidamente sobre las bibliotecas escolares. Creo que Internet es suficientemente peligroso como para que hasta cierta edad, 16 ó 18, la navegación por la Red fuera guiada, pero una vez superado ese nivel, ni los bibliotecarios, ni los documentalistas, ni cualquier otra persona a través de instituciones o de los centros que estén ofreciendo acceso público a Internet, es nadie para realizar ningún tipo de exclusión, censura o criba, exceptuando, obviamente, aquéllas que sean ilegales. Si alguien está consultando unas páginas que no son apropiadas, sencillamente está haciendo uso de su libertad. Creo que cuando se plantean esos problemas en el entorno bibliotecario, habría que actuar simplemente con un mecanismo de tiempos, porque los recursos públicos que se ponen a disposición de terceros son recursos limitados y por tanto se puede poner una limitación temporal, más o menos flexible, para la consulta, si se tienen en cuenta los contenidos de lo que se está consultando. Pero más allá, creo que no hay ninguna razón, en absoluto, para crear ningún tipo de exclusión. De hecho, los sistemas actuales de filtro funcionan tan mal que no merecen la pena, y en el caso infantil o juvenil, yo no dejaría en manos de sistemas automáticos la navegación, por parte de nuestros hijos, sino que debe haber una guía y una responsabilidad humana, que podrían ser los educadores y los bibliotecarios. Ahora más que nunca está justificada la tarea de los bibliotecarios como una guía ante la navegación en Internet.

"Lo que yo planteo es si son los criterios exclusivamente formales los que determinan el la evaluación de la Web, y si con este examen estamos resolviendo el problema. Mi respuesta es que se está haciendo mucho hincapié la forma y poco en los contenidos y que esa tendencia hay que cambiarla".



incapaz de tener una visión global de la información, a causa del desconocimiento acerca de cuáles son las herramientas que le pueden llevar directamente hacia los datos que le interesan y la pérdida de información por culpa de silencios documentales. Y en todo esto proceso los informáticos y los profesionales actuamos de manera diferente. A ellos les preocupa, sobre todo, la seguridad, cosa que para nosotros, que trabajamos en el dominio público, no es prioritario. Que algunos de los recursos lleguen a terceros no nos supone un problema, sino que se está difundiendo una labor que por definición ya es de dominio público, y en ocasiones resulta difícil hacer entender a alguien que está por debajo de mí en el escalafón que yo soy quien pone las prioridades y que para mí es prioritario un mecanismo de acceso común a varios recursos de información diferentes o que las páginas dinámicas se actualicen mejor o que el

-¿Quién debería encargarse, entonces, de orientar sobre qué información es útil o no dentro de las páginas de Internet, en caso que alguien tuviera que hacerlo?

-Hay que tener mucho cuidado con la palabra selección. La selección pura y dura sólo debe afectar a los casos extremos, es decir, la no selección de recursos que estén en idiomas desconocidos, con alfabetos diferentes, podría ser un criterio de exclusión, o la no disponibilidad de una página porque el servidor plantee problemas de entrada, o aquellos recursos que requieran un tipo de pago. En el resto de los casos, yo no sería tan crudo e iría directamente a la descripción, que puede ser muy fuerte. Es decir, de una descripción negativa, el usuario puede intuir una exclusión. Por otra parte, me preocuparía mucho que el bibliotecario o el documentalista iniciara programas de selección, convirtiéndose en una especie de censor. Creo que se puede censurar bastante describiendo,

-Quizá los padres no sepan, en general, que urge la necesidad de guiar a sus hijos delante de Internet.

-Hay un dato que siempre comento, a propósito del porcentaje de páginas pornográficas que hay en Internet. Las últimas estadísticas dicen que aproximadamente el 1,5 % de todo Internet tiene un carácter negativo, que incluye sectas, pornografía, nazismo, defensa de la anorexia y la bulimia, entre otros temas, por lo que decir que Internet está lleno de basura es absurdo, y la cifra, a todos los efectos, es ridícula. Ahora bien, lo que hay que pensar es que ese 1,5 % de entre diez mil o doce mil millones de páginas Web que hay en la actualidad es suficiente como para que el tema sea muy preocupante para la formación de nuestros hijos. Yo como adulto puedo entender que

Internet es una herramienta muy útil porque el volumen de información “indecente” -por llamarlo de alguna manera- es ridículo, pero este 1,5 % de 12.000 millones, que son centenares de millones de páginas, me puede preocupar de cara a la educación de los pequeños. Este es el doble planteamiento que yo le veo.

-¿Por qué cree usted que no se ha cumplido la profecía del E-book en la Sociedad de la Información. Qué está fallando?

-No creo que esté fallando, sino que el concepto de libro electrónico no se está explotando en los sectores más adecuados. Es decir, que las opciones que han tenido éxito tienen unas características que se acomodan bien a los mecanismos de recuperación y de navegación, como es el caso de la enciclopedia, cuyo mecanismo de navegación remite al de la estructura de Web, donde puedes saltar de un lugar a otro. En cambio, un libro tradicional llevado al formato electrónico no ofrece ningún valor añadido, respecto al soporte tradicional, ya que únicamente te ofrece la opción de la lectura lineal, y leer un documento extenso en un formato como el ordenador que, hasta la fecha, no es tan ergonómico como el papel, creo que no tiene futuro y que el libro electrónico tiene sentido en los entornos donde hay un valor añadido a la simple lectura, por ello creo que me decepcionó el formato electrónico del diccionario de la Real Academia o el “María Moliner”, que siempre había deseado tener, y que adquirí directamente en este formato. En ese sentido, cabe destacar el éxito de los métodos de aprendizaje de idiomas, con métodos interactivos, y eso mismo se podría aplicar a la formación, no sólo la reglada, sino a la promoción a la largo de la vida y la carrera profesional, y aquí el formato electrónico ofrece una de sus mayores ventajas: que el coste por unidad es ridículo, una vez que se ha generado el formato, bien para CD o para Internet. El libro electrónico que tendrá futuro será el que sepa explotar este valor añadido del dinamismo y la interactividad. Hay que darle tiempo al tiempo. Posiblemente los autores tendrán que empezar a escribir los libros de otra manera. A mí me encantaría, por ejemplo, encontrar algo así con la obra de Graham Green.

-Dentro del programa de su curso “Recuperación de Información para bibliotecarios, documentalistas y archiveros”

habla usted de Internet Invisible. ¿Qué quiere decir con este juego de palabras?

-Esto es uno de mis famosos ataques contra el googlecentrismo del que hablaba antes. Una de las afirmaciones que se ha ido transmitiendo es que, en realidad, recuperar la información en Internet es tan fácil como poner un término en Google y que por lo tanto, uno puede pensar que si aprende cómo funciona Google y lo domina va a obtener resultados válidos. El concepto de “Internet Invisible” viene a atacar esa idea, ya que Google, hoy por hoy, apenas cubre un tercio de toda la Internet que está visible, pero es que incluso hay una parte de Internet que por razones estructurales no se puede recuperar de un motor de búsqueda, y que es del orden de 2 a 50 veces mayor que la Internet visible. ¿Cuáles son, entonces, las razones por las que un motor de búsqueda no puede indizar un recurso? Pues hay muchas, pero la mayoría de ellas tienen que ver con que el recurso tiene una estructura que imposibilita su recuperación por un motor de búsqueda. Es el caso de cualquier mecanismo de base de datos, sea bibliográfica, factual o a texto completo, porque cuando Google tiene que interrogar a la base de datos no sabe qué interrogar, simplemente se le ofrece una pasarela de búsqueda y no puede proporcionarnos acceso a toda esa información que se encuentra en bases de datos. En principio, eso no sería conflictivo si no hubiera detrás de esa pasarela mucha información, pero se calcula que en la actualidad hay 40.000 catálogos de biblioteca, algunos de ellos con varios cientos de miles o millones de registros. Y por otra parte, se dice que hay de 200.000 a 500.000 bases de datos en Internet, algunas bibliográficas con 10, 15 millones de recursos, así como otras bases que pueden ser tesis completas o todas las bases de datos de libros de Amazon, que se calcula son varios millones de libros. Todo ello es lo que yo llamo Internet Invisible, como también todos los depósitos de documentos que se recuperan mediante búsqueda, es decir, la producción científica de casi diez mil revistas electrónicas, artículos que son seleccionados, evaluados y revisados. Toda esa información, invisible, no está en los motores de búsqueda, porque éstos no pueden penetrar esas pasarelas. Una información importante en cantidad y en calidad. Es lo que nosotros llamamos archipiélago de calidad. ¿Quién puede generar un catálogo de biblioteca? ¿Quién quiera?, ¿De la misma manera que cualquier persona pueda colgar una página de Internet? No, no es el caso. Un catálogo bibliográfico está hecho por profesionales que tienen una formación y una

experiencia más o menos extensa, los cuales se basan, además, en una doctrina que históricamente tiene más de un siglo, por lo tanto hay un valor añadido, porque es un producto documental muy estructurado y trabajado. ¿Qué significa todo ello? Que como media la Internet Invisible es superior en cifras que la Internet visible, y además, de mucha mayor calidad. -¿Cómo se podría, entonces, acceder a esta Internet Invisible?

-Muy sencillo, a través de herramientas específicas. La mayoría son bases de datos, así que lo que necesitamos son directorios de bases de datos, de catálogos de biblioteca, de obras de referencia (un diccionario, enciclopedia, guía telefónica), de bases de datos bibliográficas. En Estados Unidos puedes poner el nombre de una persona y te aparece el número de teléfono, cosa que aquí sería ilegal, pero allí es una información que está en bases de datos. Por tanto, es Internet Invisible, porque jamás la encontrarás a través de un motor de búsqueda.

-¿Cuál es el futuro profesional del documentalista, del bibliotecario y del archivero?

-Somos intermediarios y tenemos que actuar como tales. Se nos paga por y para eso. Los especialistas tenemos que dedicarnos a guiar, a conocer las herramientas, a dominarlas, a familiarizarnos con las nuevas que surgen y a estudiar las debilidades de algunas de ellas. Lo que no se puede pensar es que poniendo un término en un buscador se resuelven las necesidades de información de cualquier búsqueda. Eso no es tan sencillo. Nosotros somos necesarios en medio del camino y el propio mercado está (de)mostrando cuál es

nuestro futuro. En estos momentos la situación económica es ciertamente favorable y la tasa de paro que tiene la profesión es reducida, a pesar de que el crecimiento de los titulados es exponencial. Creo que lo único que me preocupa de la profesión es que hay todavía demasiado trabajo basura, por una parte, quizás, debido a la coyuntura, y en ese caso poco podemos hacer, pero también por causa, y aquí sí que podemos avanzar, del desconocimiento que tenemos de nuestras capacidades, que están subestimadas. Creo que hoy por hoy al bibliotecario o al documentalista se le asocia poco con Internet. Si lográramos incrementar la asociación del profesional con la recuperación de información en Internet, la evaluación, la descripción y la puesta a disposición de terceros de una manera organizada y estructurada, creo que nos valoraríamos y nos valorarían más. Una mejora de la imagen profesional dentro y fuera de nuestro entorno, podría ayudar a acabar con el trabajo basura.